particularly prominent. (8) Although only two peptides (glycylglycine and glutathione) have been studied in alkaline solution, it seems that some of the CH stretching and deformation lines of peptides may change as much on removal of a proton from the $-NH_3^+$ group as do those of some amino acids.[7] Some changes also occur on ionization of the carboxyl group, but no general relationship is evident.

CAMBRIDGE, MASS.

[CONTRIBUTION FROM THE BIOLOGICAL LABORATORIES, HARVARD UNIVERSITY]

Raman Spectra of Amino Acids and Related Compounds. XI. The Ionization of Cysteine[1-3]

BY DAVID GARFINKEL[4] AND JOHN T. EDSALL[5]

RECEIVED FEBRUARY 4, 1958

The Raman spectrum of cysteine has been determined at a series of *p*H values, from strongly acid solution to *p*H 11.9. The ionization of the sulfhydryl group has been followed by observing the intensity of the S–H stretching frequency near 2575, and the bending frequency near 875 cm.$^{-1}$, as functions of *p*H. At *p*H 9.4, these intensities are about half as great as in acid solution, indicating that the acidic ionization constant of the –SH group in the presence of an adjoining $-NH_3^+$ is not far from equal to that of the $-NH_3^+$ group in the presence of an adjoining –SH. This agrees with evidence from ultraviolet spectra and other data. A line at 1084 cm.$^{-1}$ appears to be associated with the un-ionized amino group. The ionization of the carboxyl group is attended by the disappearance of the line at 1739 and the appearance of strong lines at 1345 and 1400; these changes are characteristic of carboxylic acids. The loss of a proton from the $-NH_3^+$ group is attended by marked changes in the C–H stretching vibrations, similar to those already observed in glycine and β-alanine.

The Raman spectrum of the cysteine cation has been studied already.[6] However, a further study appeared highly desirable, for the ionization of each of the three acidic groups—carboxyl, ammonium and sulfhydryl—in the cysteine molecule is associated with significant changes in the spectrum. In particular, the intensity of the strong Raman line near 2570 cm.$^{-1}$, which represents a stretching frequency of the S–H bond, is an excellent indicator of the concentration of un-ionized sulfhydryl groups present. Since the sulfhydryl and ammonium groups ionize in the same *p*H range, it is important to follow spectroscopically the ionization of the individual groups in order to determine the relative concentrations of the different individual forms present.

In view of the fact that the spectra were determined photographically, the results obtained must be considered only semi-quantitative. However, in principle, this should be one of the most reliable methods available for determining the numbers of molecules in the various ionized states in aqueous solution.

Experimental

The technique for obtaining the Raman spectra has been described.[7] Solutions of cysteine (approximately 0.67 molar, in each of the various ionization states) were prepared from Mann cysteine hydrochloride (this was assayed by Dr. M. J. Hunter and found to contain at least 98% of the theoretical sulfhydryl content). In order to prevent the oxidation of cysteine to cystine, which occurs at alkaline *p*H and is known to be catalyzed by heavy metal ions,[8] all solutions were adjusted to the desired *p*H (as measured with the Beckman Model G *p*H meter) with 3 *N* KOH containing 0.01 *M* KCN (the Raman line of cyanide ion at 2081 cm.$^{-1}$ does not appear in any of our spectra, and neither does the strongest line of cystine.) The initial cysteine hydrochloride solutions were treated with charcoal and filtered through Whatman No. 4 paper. The *p*H adjustment was then made in an ice-bath. At alkaline *p*H the solutions promptly turned purple, but the color faded on standing and was reversibly bleached by light.

Results

The spectra obtained, together with the infrared spectrum of cysteine hydrochloride[9] for comparison, are presented in Table I. The assignments in this table are based in part on work by Edsall, Otvos and Rich[6] and Sheppard.[10]

Discussion

Cysteine has three observed ionization constants, the corresponding *p*K values, extrapolated to zero ionic strength, being 1.71, 8.33 and 10.78.[11] The first of these obviously is due to ionization of the carboxyl group. The assignment of the second and third is not so obvious and has caused much discussion.[12-17]

Using the notation of Hill,[18] we may set up a detailed ionization scheme for cysteine (see also ref. 17). The three ionizing groups—carboxyl (1), sulfhydryl (2) and ammonium (3)—are designated by numbers for convenience. In this no-

---

[1] The preceding paper of the series is by D. Garfinkel and J. T. Edsall, THIS JOURNAL, **80**, 3818 (1958).

[2] Taken in part from the Ph.D. thesis of David Garfinkel, Graduate School of Arts and Sciences, Harvard University, 1955.

[3] Supported by a grant from the National Science Foundation (NSF G-621) and by a special fellowship granted to one of us (J.T.E.) by the John Simon Guggenheim Memorial Foundation.

[4] Predoctoral fellow of the National Science Foundation, 1953–1954.

[5] To whom inquiries concerning this paper should be sent.

[6] J. T. Edsall, J. W. Otvos and A. Rich, THIS JOURNAL, **72**, 474 (1950).

[7] D. Garfinkel and J. T. Edsall, *ibid.*, **80**, 3807 (1958).

[8] O. Warburg and S. Sakuma, *Pflüger's Arch. ges. Physiol.*, **200**, 203 (1923).

[9] H. M. Randall, R. G. Fowler, N. Fuson and J. K. Dangl, "Infrared Determination of Organic Structures," Van Nostrand, New York, N. Y., 1949.

[10] N. Sheppard, *J. Chem. Phys.* **17**, 74, 79 (1949).

[11] H. Borsook, E. L. Ellis and H. M. Huffman, *J. Biol. Chem.*, **117**, 281 (1937).

[12] L. R. Ryklan and C. L. A. Schmidt, *Arch. Biochem.*, **5**, 89 (1944).

[13] M. Calvin in "Glutathione," edited by S. Colowick, *et al.*, Academic Press, New York, N. Y., 1954, p. 3.

[14] E. J. Cohn and J. T. Edsall, "Proteins, Amino Acids and Peptides," Reinhold Publ. Corp., New York, N. Y., 1943, Chapters 4 and 5.

[15] M. A. Grafius and J. B. Neilands, THIS JOURNAL, **77**, 3389 (1955).

[16] R. E. Benesch and R. Benesch, *ibid.*, **77**, 5877 (1955).

[17] J. T. Edsall and J. Wyman, "Biophysical Chemistry," Vol. I, Academic Press, New York, N. Y., 1958, Chapter 9.

[18] T. L. Hill, *J. Phys. Chem.*, **48**, 101 (1944).

TABLE I[a]

RAMAN SPECTRA OF CYSTEINE

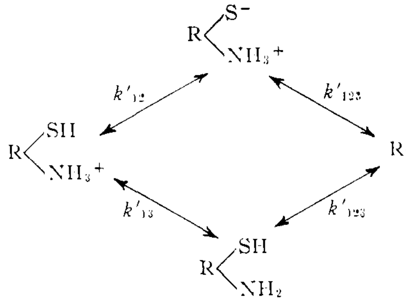
Hydrochloride	Isoelectric pH 5.75 and 7.42	K Salt pH 9.36	di-K Salt pH 11.93	Infrared	Assignments
362(vw)					
	398(vw)				
		417(vw)			
		432(vw)			
	465(vw)				
480(vw)	483(w)				
513(vw)					
529(w)	535(w)	529(1vb)	536(w)		Skeletal deformation??
			610(vw)		
614(w)	621(1vb)	630(w)	631(w)		Skeletal deformation??
	654(vw)		657(vw)		
679(6b)	683(6b)	689(6b)	682(6b)		C–S stretching
			703(w)		
	716(vw)	713(vw)			
743(w)			742(w)		
764(3b)	757(vw)		762(vw)		
778(1b)	777(2b)				
		788(2)	793(w)		
800(2b)					
	816(2b)	819(w)			
		831(2b)	833(2)		
864(3vb)	874(3b)	880(vw)			SH deformation
		901(2b)	908(4b)		
935(3b)	931(3b)		935(vw)		
		950(1b)	953(w)		
	966(vw)		977(w)		
994(2b)	990(1vb)	995(vw)			
			1028(w)		
		1045(2b)	1038(vw)		
1068(2)	1057(1vb)		1062(vw)		
		1084(w)	1085(3)		NH_2 motion
				1109(m)	
1118(w)	1121(1vb)	1129(w)			
1152(1b)				1145(m)	
	1166(vw)				
1213(2)	1208(3)	1213(1)	1205(1b)	1206(s)	Skeletal stretching??
				1223(s)	(COH deformation in COOH)
			1235(w)		
		1250(1b)			
1267(1vb)	1269(w)				NH_3^+ deformation?
			1282(w)		
	1306(2)	1297(4)	1302(2)		
				1317(w)	
1326(1)		1332(w)			
	1345(5)	1345(5b)	1345(6b)		COO^- motion
				1353(m)	
1364(w)					
			1385(2b)		
	1397(4)	1401(4)	1398(3)	1399(m)	COO^- symm. stretching
			1410(2b)		
1423(4)	1426(5)	1432(5)	1428(6)		CH_2 deformation
			1473(vw)		
				1493(w)	
				1517(m)	NH_3^+ deformation
				1574(w)	
				1655(w)	NH_3^+ deformation?
1739(3b)				1742(s)	CO stretching
				1984(w)	
				2087(w)	
2539(w,vb)				2544(w)	H-bonded SH
2577(7)	2578(7b)	2581(4b)			SH stretching
			2780(vw)		
		2848(1)	2844(vw)		

	2856(vw)	2858(w)			
2878(vw)				2881(s)	(CH stretching)
	2924(1b)	2916(5)	2907(5b)		CH stretching generally
2945(6)	2955(6)	2951(8b)	2946(2b)		
2961(8)	2963(7)				
	2996(4)				
3005(5)	3010(4)	3007(vw)			
				3378(w)	(NH$_3^+$ stretching?)
				3571(w)	

[a] Intensities are reported by numbers 1, 2, ..., for lines of increasing intensity, except that the weakest lines are designated weak (w) or very weak (vw). Infrared intensities (from ref. 9) are reported as weak (w), medium (m), or strong (s). Broad lines are designated (b), very broad lines (vb).

tation the ionization of a proton from a specific group of a particular microscopic form is described by a dissociation constant with a subscript in which the last digit denotes the group from which a proton is being removed, while the preceding digits indicate the groups from which protons have already been removed. Thus k_{123} represents the dissociation of a proton from the ammonium group of a molecule from which the ionizable carboxyl and sulfhydryl protons already have been removed.

The carboxyl group of cysteine may safely be assumed to be completely ionized about pH 6. We may then set up an ionization scheme for the ammonium and sulfhydryl groups



Here R denotes the rest of the molecule, including the ionized carboxyl group. Hence the net charge on each indicated species is one unit more negative than the formula as written would suggest.

The macroscopic constants, K'_2 and K'_3, as obtained from a regular pH titration, in which pH is determined as a function of the number of protons removed from the isoelectric molecule, are given by

$$K'_2 = (\mathrm{H}^+)\left[\left(\mathrm{R}\langle^{\mathrm{S}^-}_{\mathrm{NH_3^+}}\right) + \left(\mathrm{R}\langle^{\mathrm{SH}}_{\mathrm{NH_2}}\right)\right]\Big/\left(\mathrm{R}\langle^{\mathrm{SH}}_{\mathrm{NH_3^+}}\right) = k'_{12} + k'_{13} \tag{1}$$

$$K'_3 = (\mathrm{H}^+)\left[\left(\mathrm{R}\langle^{\mathrm{S}^-}_{\mathrm{NH_2}}\right)\Big/\left[\left(\mathrm{R}\langle^{\mathrm{S}^-}_{\mathrm{NH_3^+}}\right) + \left(\mathrm{R}\langle^{\mathrm{SH}}_{\mathrm{NH_2}}\right)\right]\right] \tag{2}$$

$$1/K'_3 = 1/k'_{123} + 1/k'_{132} \tag{3}$$

Here (H$^+$) denotes the antilogarithm of the measured pH, and the other symbols denote the *concentrations* of the species involved. We note that K'_3, in particular, should be considerably larger than the limiting value of K_3 at zero ionic strength,[11] which already has been given as 10.78. The species (–S–R–NH$_2$) carries a double negative charge; its two conjugate acids each carry a single negative charge. At ionic strength 0.16, it may be calculated roughly, using a simplified form of the Debye–Hückel equations (ref. 17, pp. 441–445), that pK'_3 should be near 10.4. At the considerably higher ionic strengths employed here, the exact value of pK'_3 is difficult to calculate but should not be very far from 10.4. The value of pK'_2 should be much less sensitive to ionic strength and fairly close to the limiting value for pK_2 of 8.33 at zero ionic strength.

The microscopic constants k'_{12}, k'_{13}, k'_{123}, k'_{123} may be approximately evaluated by studying the ionization of substances with ionizable groups similar to those of cysteine, but with certain groups blocked so that they cannot react. This may be done by converting an acidic group –RH to an alkyl derivative –RR'; it is assumed that –RH and –RR' are equivalent in their action on the other acidic groups in the molecule, and there is much evidence for this.[14,17,19] However, this assumption can be considered an approximation only, since charge effects and steric effects may be changed by the substitution.

Studies by this method have been carried out by Ryklan and Schmidt,[12] who studied S-ethylcysteine [$\mathrm{C_2H_5SCH_2CH(COO^-)NH_3^+}$] and effectively found a value for pk_{13} of 8.60. Likewise Grafius and Neilands[15] studied cysteine betaine [$\mathrm{HSCH_2CH(COO^-)N(CH_3)_3^+}$] and found a value corresponding to pk_{12} of 8.65.

Alternatively one may determine the concentration of a specific ionized form or sum of forms (for instance, the sum of all forms with the sulfhydryl group ionized or un-ionized) by physical methods. Benesch and Benesch[16] used ultraviolet spectrophotometry in the region 230–240 mμ to determine total concentration of ionized sulfhydryl groups. This requires making the assumption that the extinction coefficients of all the forms containing ionized sulfhydryl groups are the same, which is probably a good approximation but is not likely to be exactly true. From their work the values of pk_{12} and pk_{13} are 8.53 and 8.86, respectively.[19a]

In this study we have made analogous preliminary estimates based on the intensity of the

---

(19) H. C. Brown, D. H. McDaniel and O. Häfliger in E. A. Braude and F. E. Nachod, "Determination of Organic Structures by Physical Methods," Academic Press, New York, N. Y., 1955, Chapter 14.

(19a) For a more detailed discussion of the evaluation of the microscopic pK values, in such substances as cysteine and tyrosine, see J. T. Edsall, R. B. Martin, and B. R. Hollingworth, *Proc. Nat. Acad. Sci.*, (U.S.), **44**, June (1958).

Raman scattering due to the –SH stretching motion near 2575 cm.$^{-1}$ and the –SH bending motion at 864–880 cm.$^{-1}$. Both lines disappear as the sulfhydryl group ionizes. The line near 2575 is very intense in the Raman spectrum, although far weaker in the infrared. We may compare the relative intensities of these lines, at a series of *p*H values, with the intensity of the strong C–S stretching frequency at 680–690 cm.$^{-1}$, which may be expected to be nearly unaffected by *p*H, and is indeed extremely strong at all *p*H values from 1 to 12 (Table I). In contrast, the 2575 line is very intense in acid solution, where the cysteine cation is present, or at *p*H 5.7–7.4, in a range where the isoelectric molecule is present; but at *p*H 9.36 its intensity has fallen to approximately half this value, and at *p*H 11.93 it has completely disappeared. The –SH bending frequency near 870 cm.$^{-1}$ behaves in parallel fashion.

This indicates that at *p*H 9.36, nearly midway between pK'_2 and pK'_3, approximately half the sulfhydryl groups in the solution are ionized. In order to satisfy equations 1, 2 and 3, this implies that the ratio of k'_{12} to k'_{13} is not far from unity. If $k'_{12} >> k'_{13}$, the intensity of the 2575 line should have fallen nearly to zero at *p*H 9.36; if $k'_{12} << k'_{13}$ the intensity of this line should have been practically as great as at low *p*H values. Both of these extreme hypotheses had been suggested[13,14]; our findings agree with those of other workers cited earlier in this paper in ruling out such conclusions. A more quantitative study of the ionization of the sulfhydryl group, by the measurement of Raman line intensities as a function of *p*H, is to be undertaken shortly, when an instrument for accurate intensity recordings by photoelectric methods becomes available to us.

It should also be possible to follow the individual ionization steps by measuring the intensity of the frequencies associated with the amino group as a function of *p*H. The N–H stretching frequencies near 3305 and 3370, characteristic of the uncharged amino group,[20–22] were obscured in our studies, partly because they lie at the edge of the very intense water band, partly because the continuous background generally becomes more intense in alkaline solutions, which are relatively difficult to clarify. These difficulties should be much diminished by photoelectric recording.

One line, at 1084 cm.$^{-1}$, is non-existent in solutions of the cation and dipolar ion, weak at *p*H 9.36, and quite strong at *p*H 11.93. This behavior suggests that it is associated with the uncharged amino group, but we make no attempt at this time to assign it to a definite vibrational mode.

The presence of the very broad line at 2539 cm.$^{-1}$ suggests that in acid solutions of cysteine there may be appreciable hydrogen bonding between sulfhydryl groups and water molecules (compare Benesch).[23]

The ionization of the carboxyl group may be seen from Table I to be accompanied by the usual characteristic changes[6,22,24]—the disappearance of the C=O stretching frequency near 1740, and the appearance of strong lines at 1400 and 1345.

In cysteine, as in glycine, the bending frequency of the methylene group is low, lying between 1423 and 1432, and showing no apparent relation to the state of ionization of the molecule. On the other hand the C–H stretching frequencies appear to be markedly affected by ionization of the amino group. (The ionization of the sulfhydryl group may of course play a part also.) The lines at 2963 and 2996, which are strong in acid solution, fade out at high *p*H, while the lines between 2900 and 2950 remain strong and probably become broader. The increasing strength of the background, generally observed in alkaline solutions, makes precise observation of these effects difficult. Nevertheless their general character appears clear, and the changes found in alkaline solution are strongly reminiscent of those previously observed for the C–H stretching lines of glycine and β-alanine in alkaline solution.[22]

CAMBRIDGE, MASS.

---

[20] J. T. Edsall, *J. Chem. Phys.*, **5**, 225 (1937).

[21] J. T. Edsall, *ibid.*, **8**, 520 (1940).

[22] M. Takeda, R. E. S. Iavazzo, D. Garfinkel, I. H. Scheinberg and J. T. Edsall, THIS JOURNAL, **80**, 3813 (1958).

[23] R. Benesch, in "Glutathione," edited by S. Colowick, *et al.*, Academic Press, New York, N. Y., 1954, p. 31.

[24] J. T. Edsall, *J. Chem. Phys.*, **5**, 508 (1937).